

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-122508

(P2003-122508A)

(43) 公開日 平成15年4月25日 (2003.4.25)

(51) Int.Cl.⁷

G 0 6 F 3/06

識別記号

3 0 1

F I

G 0 6 F 3/06

テーマコード* (参考)

3 0 1 J 5 B 0 6 5

審査請求 未請求 請求項の数10 O L (全 17 頁)

(21) 出願番号 特願2001-316860 (P2001-316860)

(22) 出願日 平成13年10月15日 (2001. 10. 15)

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 副島 健一

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

(72) 発明者 浅野 正靖

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

(74) 代理人 100068504

弁理士 小川 勝男 (外2名)

最終頁に続く

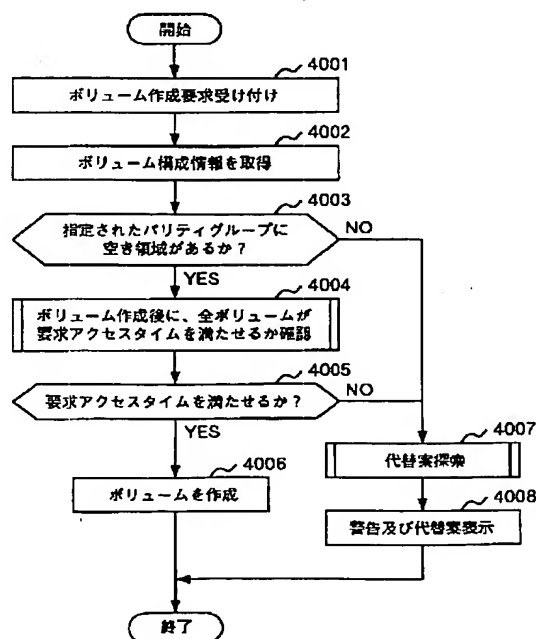
(54) 【発明の名称】 ボリューム管理方法及び装置

(57) 【要約】

【課題】 複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームを作成する方法において、各ボリュームの要求性能を考慮した作成を行う。

【解決手段】 記憶容量と要求平均性能情報を指定するボリューム作成要求を受け付け、複数台の記憶装置に亘って要求された記憶容量を満足する空き領域があるか否か判定し、空き領域が存在する場合に作成要求されたボリュームと記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して既存ボリュームに作成要求されたボリュームを加えた全ボリュームの各々がその要求平均性能を満足するか否か判定し、全ボリュームが各要求平均性能を満足する場合に、作成要求されたボリュームを作成する。

図 4



【特許請求の範囲】

【請求項1】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームを設定する方法において、記憶容量と要求平均性能情報を指定するボリューム作成要求を受け付けるステップと、複数台の前記記憶装置に亘って要求された前記記憶容量を満足する空き領域があるか否かを判定するステップと、前記空き領域が存在する場合に作成要求されたボリュームと前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記作成要求されたボリュームを加えた全ボリュームの各々がその要求平均性能を満足するか否かを判定するステップと、前記全ボリュームが各要求平均性能を満足する場合に、前記作成要求されたボリュームを設定するステップとを有することを特徴とするボリューム管理方法。

【請求項2】前記全ボリュームが同一のパーティグループに属するボリュームであり、前記作成要求されたボリュームは、異なるパーティグループからのボリューム移動要求であることを特徴とする請求項1記載のボリューム管理方法。

【請求項3】指定された前記記憶容量と指定された前記要求平均性能のうち少なくとも一方の条件を満足する空き領域が存在しない場合に、満足しない条件について所定割合の削減又は増加を加えた条件について該当する前記判定するステップを再実行することを特徴とする請求項1記載のボリューム管理方法。

【請求項4】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームが設定されるストレージ中に新規ボリュームのための空き領域を探索する方法であって、要求平均性能情報を指定する空き領域の探索要求を受け付けるステップと、各空き領域と前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記新規ボリュームを加えるとした場合に全ボリュームの各々がその要求平均性能を満足するか否かを判定するステップと、前記の判定条件を満足するような空き領域についての情報を表示するステップとを有することを特徴とするボリューム管理方法。

【請求項5】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームを設定するボリューム管理装置において、記憶容量と要求平均性能情報を指定するボリューム作成要求を受け付ける手段と、複数台の前記記憶装置に亘って要求された前記記憶容量を満足する空き領域があるか否かを判定する手段と、前記空き領域が存在する場合に作成要求されたボリュームと前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記作成要求されたボリュームを加えた全ボリュームの各々がその要求平均性能を満足するか否かを判定する手段と、前記全ボリュームが各要求平均性能を満足する場合に、

前記作成要求されたボリュームを設定する手段とを有することを特徴とするボリューム管理装置。

【請求項6】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームが設定されるストレージ中に新規ボリュームのための空き領域を探索するボリューム管理装置であって、要求平均性能情報を指定する空き領域の探索要求を受け付ける手段と、各空き領域と前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記新規ボリュームを加えるとした場合に全ボリュームの各々がその要求平均性能を満足するか否かを判定する手段と、前記の判定条件を満足するような空き領域についての情報を表示する手段とを有することを特徴とするボリューム管理装置。

【請求項7】コンピュータに複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームの設定機能を実現させるためのプログラムにおいて、コンピュータに、記憶容量と要求平均性能情報を指定するボリューム作成要求を受け付ける機能、複数台の前記記憶装置に亘って要求された前記記憶容量を満足する空き領域があるか否かを判定する機能、前記空き領域が存在する場合に作成要求されたボリュームと前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記作成要求されたボリュームを加えた全ボリュームの各々がその要求平均性能を満足するか否かを判定する機能、および前記全ボリュームが各要求平均性能を満足する場合に、前記作成要求されたボリュームを設定する機能を実現させるためのプログラム。

【請求項8】コンピュータに複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームが設定されるストレージ中に新規ボリュームのための空き領域探索機能を実現させるためのプログラムであって、コンピュータに、要求平均性能情報を指定する空き領域の探索要求を受け付ける機能、各空き領域と前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記新規ボリュームを加えるとした場合に全ボリュームの各々がその要求平均性能を満足するか否かを判定する機能、および前記の判定条件を満足するような空き領域についての情報を表示する機能を実現させるためのプログラム。

【請求項9】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームを設定するストレージと、前記ストレージの情報にアクセスするクライアント計算機と、前記ストレージ内のボリュームを管理する管理用計算機と、前記ストレージ、前記クライアント計算機及び前記管理用計算機間を接続するネットワークとを有するコンピュータシステムの前記管理用計算機によるボリューム設定方法において、記憶容量と要求平均性能情報を指定するボリューム作成

要求を受け付けるステップと、複数台の前記記憶装置に亘って要求された前記記憶容量を満足する空き領域があるか否か判定するステップと、前記空き領域が存在する場合に作成要求されたボリュームと前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記作成要求されたボリュームを加えた全ボリュームの各々がその要求平均性能を満足するか否か判定するステップと、前記全ボリュームが各要求平均性能を満足する場合に、前記作成要求されたボリュームを設定するステップとを有することを特徴とするボリューム管理方法。

【請求項10】複数台の物理的な記憶装置に亘って少なくとも1つの論理的なボリュームを設定するストレージと、前記ストレージの情報にアクセスするクライアント計算機と、前記ストレージ内のボリュームを管理する管理用計算機と、前記ストレージ、前記クライアント計算機及び前記管理用計算機間を接続するネットワークとを有するコンピュータシステムの前記ストレージ中に新規ボリュームのための空き領域を探索する前記管理用計算機による方法であって、

要求平均性能情報を指定する空き領域の探索要求を受け付けるステップと、各空き領域と前記記憶装置を共用する既存ボリュームについての記憶手段上の各要求平均性能情報を参照して前記既存ボリュームに前記新規ボリュームを加えるとした場合に全ボリュームの各々がその要求平均性能を満足するか否か判定するステップと、前記の判定条件を満足するような空き領域についての情報を表示するステップとを有することを特徴とするボリューム管理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ストレージ内の物理的記憶装置上に定義されるボリュームの管理方式に係わり、特に他のボリュームの性能に与える影響を考慮しながらボリュームの作成や移動を行なうための処理方式に関する。

【0002】

【従来の技術】2台以上の物理ディスクをグループ化して冗長性をもたせ、信頼性や性能を向上させるRAID (Redundant Array of Inexpensive Disks) と呼ばれる技術がある。RAID技術を採用しているストレージを使用する場合、まずストレージ内の物理ディスク（物理記憶装置）を2台以上集めてパリティグループと呼ばれるグループを構成する。そしてパリティグループ内にボリュームと呼ばれる論理記憶領域を定義し、計算機側からはこのボリュームを使用する。多くの場合、1個のパリティグループ内に2個以上のボリュームが定義可能である。定義例を図1(a)に示す。図1(a)では、物理ディスク1001、物理ディスク1002、物理ディスク1003及

び物理ディスク1004が1個のパリティグループ1005を構成し、その中にボリューム1006、ボリューム1007及びボリューム1008が定義されている。このときパリティグループ1005内に定義されているボリューム1006、ボリューム1007及びボリューム1008は、全て物理ディスク1001、物理ディスク1002、物理ディスク1003及び物理ディスク1004を使用している。この異なるボリュームが同じ物理ディスクを共有しているという状態は、RAID技術を採用しているストレージを使用する場合だけでなく、単一の物理ディスク内に2個以上のボリュームを定義する場合にも存在する。

【0003】このようなRAID技術を使用して、複数の物理ディスクからパリティグループを構成する方法は数通り定義されており、これらの基準はRAIDレベルと呼ばれる。例えば複数の物理ディスクに一定サイズ毎に順にデータを書くRAIDレベル0（ストライプとも呼ばれる）や、2個の物理ディスクの両方に読み書きを行なうRAIDレベル1（ミラーとも呼ばれる）、複数の物理ディスクの1台を除きデータを分散して記憶させそのデータを記憶させなかった物理ディスクには、他の物理ディスクが故障して使用不可能になっても、当該データを復活できるような冗長データを記述するRAIDレベル4やRAIDレベル5（これらは冗長データの持ち方が異なる）等がある。

【0004】物理ディスク全体の集合ではなく、物理ディスクの一部領域の集合を使用してパリティグループを構成することが可能なストレージがある。この場合、1個の物理ディスクは複数のパリティグループに属することができる。例を図1(b)に示す。パリティグループ1015は、物理ディスク1009、物理ディスク1010、物理ディスク1011及び物理ディスク1012のそれぞれの一部領域から構成されており、当該パリティグループ内にはボリューム1016が定義されている。パリティグループ1013は、物理ディスク1009、物理ディスク1010のそれぞれ一部領域から構成されており、当該パリティグループ内にはボリューム1017が定義されている。ここで物理ディスク1009、物理ディスク1010、物理ディスク1011及び物理ディスク1012は、それぞれ2個のパリティグループに属している。

【0005】RAID技術を採用しているストレージの幾つかは、パリティグループ内に直接ボリュームを定義するのではなく、パリティグループ内に論理ディスクを定義し、その後論理ディスクの分割や結合を数階層に渡って繰り返し、最上層の論理ディスクをボリュームとして定義する機能を持つ。例えば図1(b)において、パ

リティグループ1013内とパリティグループ1014内にそれぞれ論理ディスクを定義し、それらを結合したものをボリュームとして見せる等である。

【0006】以上で述べたRAIDに関する文献として、Mark Farley, "Building Storage Networks", Network Professional's Library, Osborne がある。

【0007】一方、近年、複数のストレージとそれらのストレージを使用する複数のクライアント計算機をファイバチャネル等の高速なネットワークで接続したSAN (Storage Area Network) が注目を集めている。SAN環境では、1台のクライアント計算機が2台以上のストレージを同時に使用可能であるとともに、1台のストレージを2台以上のクライアント計算機から使用可能である。SANの構成例を図2に示す。図2では、ストレージ2005、ストレージ2006、ストレージ2007と、これらストレージを利用するクライアント計算機2002、クライアント計算機2003と、これらストレージを管理する管理計算機2004が全てネットワーク2001により接続されている。ここで管理計算機2004は、ネットワーク2001に接続されているストレージ内の構成定義やアクセス許可制御等を行なうための管理用計算機である。図2に示したSAN環境において、管理計算機2004からアクセス許可さえ与えられれば、クライアント計算機2002は、ストレージ2005、ストレージ2006及びストレージ2007の全てを使用可能である。同様にクライアント計算機2003とクライアント計算機2004がストレージ2005を共有することも可能である。

【0008】図2におけるストレージ2005、ストレージ2006及びストレージ2007のようなSAN環境で使用するストレージは、前述のRAID技術を採用していることが多い。このときストレージ2005中に同じ物理ディスクを使用した複数のボリュームを作成し、それぞれのボリュームをクライアント計算機2002とクライアント計算機2003が使用することも可能である。

【0009】以上で述べたSANに関する文献として、「データストレージ・レポート2000」、日経コンピュータ別冊、2000年、がある。

【0010】パリティグループやボリュームの作成・移動作業には、ストレージ管理プログラムを使用する。このストレージ管理プログラムは、SANに接続された管理計算機やストレージ備え付けの端末等で動作する。このとき処理対象ストレージ内の物理ディスク容量や定義テーブル等のストレージ内の必要な資源が全て利用可能であれば、要求された操作は常に実行される。

【0011】一方、ストレージの管理プログラムには、物理ディスクやボリュームの性能管理に関する機能を持

つものがある。ある管理プログラムは、物理ディスク内のボリュームの平均アクセスタイムを監視し、予め指定された平均アクセスタイムの上限以上になったら警告する機能を持つ。また別の管理プログラムは、性能履歴を蓄積しておき、後にこの情報を利用してボリュームの再配置による性能最適化を行なう機能を持つ。

【0012】

【発明が解決しようとする課題】図1(a)、(b)に示すように、2個以上のボリュームが同じ物理ディスクを使用している場合、あるボリュームにアクセスが集中すると、同じ物理ディスクを使用している他のボリュームの性能が低下する可能性がある。例えば図1(a)においてボリューム1006にアクセスが集中すると、同じ物理ディスクを使用しているボリューム1007およびボリューム1008の性能が低下する等である。別の見方をする、既存ボリュームの性能が要求されている性能を出している状態で、当該ボリュームと同じ物理ディスクを使用して新たなボリュームを作成したり、他の物理ディスクからボリュームを移動してきたりすると、既存ボリュームの性能が要求されている性能を満たさなくなる可能性がある。例えば図1(a)において、ボリューム1006の現在の平均アクセスタイムが18ミリ秒であり、要求されている平均アクセスタイム20ミリ秒を満たしているとき、パリティグループ1005内に新たなボリュームを作成すると、ボリューム1006の平均アクセスタイムが23ミリ秒に低下し、要求されている平均アクセスタイム20ミリ秒を満たさなくなる等である。

【0013】図2で示したSAN環境のように、1個の物理ディスクを複数のクライアント計算機から利用することができる環境では、上述した問題は特に注意する必要がある。これは異なるクライアント計算機から利用されるボリュームが、同じ物理ディスク上に定義することができるためである。このような場合、あるクライアント計算機が利用するために新たにボリュームを作成すると、別のクライアント計算機が使用しているボリュームの性能が低下するということが起きる。さらにSAN環境で使用されているストレージの多くはRAID技術を採用しており、上記のように物理ディスクとボリュームとの関係は多階層に渡る分割・結合を繰り返し複雑になっている。このような環境では、どのボリュームとどのボリュームが同じ物理ディスクを使用しているかを管理者が手作業で確認しながらボリューム構成の変更を行なうことは非現実的である。

【0014】既存のストレージが以上で述べた問題を持つものに対し、従来の管理プログラムやストレージでは、他のボリュームの性能に与える影響を考慮しながらボリュームの作成・移動を行なう機能は提供されていない。

【0015】さらに従来の性能に関する機能を持つ管理プログラムは、ボリュームの作成・移動時には他のボリ

ュームの性能に対する影響は考慮せず、実際にポリリュームを使用してから性能低下の検出及び改善を図る方式をとっている。そのため各ポリリュームが一時的に要求される平均性能を満たせなくなったり、性能低下を回避するポリリューム再配置を行なった後でも再び別のポリリュームの性能が低下したりする可能性がある。

【0016】本発明の目的は、各ポリリュームの要求性能を考慮したポリリュームの作成・移動操作を行なう管理方式を提供することにある。

【0017】

【課題を解決するための手段】本発明は、複数台の物理的な記憶装置に亘って少なくとも1つの論理的なポリリュームを設定する方式において、記憶容量と要求平均性能情報を指定するポリリューム作成要求を受け付け、複数台の記憶装置に亘って要求された記憶容量を満足する空き領域があるか否か判定し、空き領域が存在する場合に作成要求されたポリリュームと記憶装置を共用する既存ポリリュームについての記憶手段上の各要求平均性能情報を参照して既存ポリリュームに作成要求されたポリリュームを加えた全ポリリュームの各々がその要求平均性能を満足するか否か判定し、その全ポリリュームが各要求平均性能を満足する場合に、作成要求されたポリリュームを設定するポリリューム管理方式を特徴とする。

【0018】また本発明は、複数台の物理的な記憶装置に亘って少なくとも1つの論理的なポリリュームが設定されるストレージ中に新規ポリリュームのための空き領域を探索する方式であって、要求平均性能情報を指定する空き領域の探索要求を受け付け、各空き領域と記憶装置を共用する既存ポリリュームについての記憶手段上の各要求平均性能情報を参照して既存ポリリュームにこの新規ポリリュームを加えるとした場合に全ポリリュームの各々がその要求平均性能を満足するか否か判定し、この判定条件を満足するような空き領域についての情報を表示するポリリューム管理方式を特徴とする。

【0019】

【発明の実施の形態】第1の実施例として、ポリリュームを作成する際に、他ポリリュームの平均アクセスタイム（以下アクセスタイム）が当該ポリリュームに要求される平均アクセスタイム（以下要求アクセスタイム）より大きくなならないことを確認し、その可能性があれば警告と代替案を表示する手順を示す。なおアクセスタイムは、値が小さいほうが短時間でデータにアクセスでき高性能であるものとする。また以後（要求）アクセスタイムを満たすとは、（要求）アクセスタイム以下のアクセスタイムで平均的にアクセス可能であることを意味する。

【0020】図3は、図2に示す管理計算機2004の構成を示すブロック図である。ここでストレージ3011は、図2のストレージ2005、ストレージ2006、ストレージ2007のいずれか1個に該当する。またストレージ3011内には、物理ディスク3015と

物理ディスク3016が存在する。ポリリューム作成要求受け付け部3001は、管理者が発行したポリリューム作成要求と、作成するポリリュームの容量、その作成先パーティグループの情報、当該ポリリュームへの平均入出力要求数（以下、平均I/O数）及び要求アクセスタイムを受け取る。ポリリューム作成要求処理部3003は、ポリリューム作成要求受け付け部3001が受け取った要求を受け取り、要求アクセスタイム維持確認部3005、ポリリューム構成情報取得部3008、代替案探索部3004、ポリリューム作成要求発行部3007及び警告・代替案表示部3002と連携しながら処理を行なう。

【0021】ポリリューム構成情報取得部3008は、物理ディスク3015および物理ディスク3016上に定義されているポリリュームの構成情報取得要求をストレージ3011内のポリリューム構成情報提供部3013に発行し、結果を取得する。ポリリューム作成要求発行部3007は、ストレージ3011内のポリリューム構成要求変更部3012にポリリューム作成要求を発行し、ストレージ3011内の物理ディスク3015および物理ディスク3016上に定義されているポリリュームの構成情報を変更する。ストレージ3011内のポリリューム構成提供部3013は、ポリリューム構成保持部3014が保持するポリリュームの構成情報を参照し、ポリリューム構成情報取得部3008に提供する。ポリリューム構成情報変更部3012は、ポリリューム構成情報保持部3014が保持するポリリュームの構成情報をポリリューム作成要求発行部3007からのポリリューム作成要求に従って変更する。ポリリューム構成情報保持部3014は、ストレージ3011内の物理ディスク3015および物理ディスク3016上に定義されているポリリュームの構成情報を保持し、ポリリューム構成情報提供部3013やポリリューム構成情報変更部3012からの要求にしたがい、保持するポリリューム構成情報の提供・変更を行なう。要求アクセスタイム維持確認部3005は、ポリリューム構成情報取得部3008、要求アクセスタイム保持部3006、ポリリューム作成後アクセスタイム決定部3009及び代替案探索部3004と連携しながら、ポリリューム作成要求処理部3003からの要求に従い、指定されたポリリュームを作成した場合に他のポリリュームのアクセスタイムが要求アクセスタイムを満たせるかどうかを確認する。要求アクセスタイム保持部3006は、ポリリュームごとに要求される性能情報として、要求アクセスタイムを保持しており、要求アクセスタイム維持確認部3005からの要求に従い、指定されたポリリュームの要求アクセスタイムを提供する。ポリリューム作成後アクセスタイム決定部3009は、要求アクセスタイム維持確認部3005からの要求に従い、指定された位置にポリリュームを作成した場合に各ポリリュームが満たすことのできるアクセスタイムを決定する。代替案探索部3004は、要求アクセスタイム維持確認部3005からの要求に従い、指定

された容量の指定されたアクセスタイムを満たす空き領域を探索する。警告・代替案表示部3002は、ボリューム作成要求処理部3003からの要求に従い、利用者から要求されたボリューム作成操作を行なうと他ボリュームのアクセスタイムが当該ボリュームの要求アクセスタイムを満たさなくなる旨を表示するとともに、代替案探索部3004で探索した代替案を表示する。

【0022】本実施例では、平均利用情報の例として、平均入出力要求数を使用しており、入出力データサイズは固定を仮定している。本実施例では図3に示したように、発明の内容を全て管理計算機2004内で実施しているが、任意の処理ブロックもしくは任意の処理ブロックを分割したブロックの一部を、ストレージ3011や図2に示すクライアント計算機2002やクライアント計算機2003、もしくは管理計算機2004上に配置してもよい。

【0023】図4は、図2の管理計算機2004内で、本発明を使用してボリューム作成要求を処理する際の処理の流れを示している。まずステップ4001で、ボリューム作成要求と作成するボリュームの容量、当該ボリュームの平均I/O数、要求アクセスタイム及びボリューム作成先パリティグループの情報を受け取る。この処理は、図3におけるボリューム作成要求受け付け部3001で行なう。なお当該ステップにおいて、ボリューム作成要求と共に要求アクセスタイムや平均I/O数が指定されなければ、以後の処理では予め決めておいた固定値、例えば50msecと1000要求/秒を使用する。ステップ4001の次にステップ4002で、図3のボリューム構成情報取得部3008によって上記のボリューム作成先パリティグループを構成する全ての物理ディスクを使用している全てのパリティグループの構成情報を取得する。さらにステップ4003で、作成要求されたボリュームの容量が指定されたボリューム作成先パリティグループに存在するかを調べる。もし十分な空き領域がない場合は、ステップ4007で、指定された容量、指定された要求アクセスタイムを満たす空き領域を探索し、ステップ4008で、結果を警告とともに代替案として表示し終了する。ここで空き領域有無判断ステップ4003は図3のボリューム作成要求処理部3003で、代替案探索ステップ4007は代替案探索部3004で、警告及び代替案表示ステップ4008は警告・代替案表示部3002で処理する。

【0024】一方、ステップ4003で指定されたボリューム作成先パリティグループに指定された容量の空き容量があると判断すれば、ステップ4004で、要求されたボリュームを作成した後に、指定されたパリティグループを構成する物理ディスク上に定義されている全てのボリュームのアクセスタイムが、要求アクセスタイムを満たせるかを確認する。この処理は、図3の要求アクセスタイム維持確認部3005で行う。この確認の結

果、ステップ4005で要求アクセスタイムを満たせると判断すれば、ステップ4006でボリューム作成要求をストレージに送る。このときステップ4005は図3のボリューム作成要求処理部3003で、ステップ4006はボリューム作成要求発行部3007で処理される。ステップ4005の判断の結果、ボリューム作成後、1個以上のボリュームのアクセスタイムが要求アクセスタイムを満たせなければ、ステップ4007で代替案を探索し、ステップ4008で前記探索結果を警告と共に表示する。ここでステップ4007は代替案探索部3004で、ステップ4008は警告・代替案表示部3002で処理される。

【0025】図5は、図4に示すボリューム作成処理後に全ボリュームが要求アクセスタイムを満たせるかを確認するステップ4004の処理手順を示している。まずステップ5001で、同じパリティグループに属する既存ボリュームと新たに作成するボリュームの一覧を作成する。ただし同一の物理ディスクを使用している他のパリティグループがあれば、そのグループ内のボリュームも一覧に含める。次にステップ5002で、このボリューム一覧に含まれる全てのボリュームの要求アクセスタイムと平均I/O数を取得する。次にステップ5003で、要求されたボリューム作成後の選択したボリュームに割り当てるアクセスタイムを決定する。次にステップ5004で、ボリューム一覧の中に要求アクセスタイム維持確認を行っていないボリュームがあるかどうかを調べ、もしなければステップ5008で全てのボリュームが要求アクセスタイムを満たせると判断して終了する。しかしステップ5004で、要求アクセスタイム維持確認を行っていないボリュームがまだあると判断すれば、ステップ5005でそのうちの1個を選択する。次にステップ5006で、先に決定したアクセスタイムがステップ5002で取得した要求アクセスタイムを満たすかどうかを判断する。それが成り立てばステップ5004に戻って上記処理を繰り返す。成り立たなければ、ステップ5007で要求アクセスタイムを満たせないボリュームが存在すると判断して終了する。

【0026】以上で述べたステップのうち、ステップ5001の既存ボリュームの情報は図3のボリューム構成情報取得部3008で取得し、またステップ5002の要求アクセスタイムは要求アクセスタイム保持部3006から取得する。またボリューム作成後アクセスタイム決定ステップ5003は、図3のボリューム作成後アクセスタイム決定部3009で処理し、それ以外のステップは、全て要求アクセスタイム維持確認部3005で処理する。

【0027】本実施例では、ステップ5001で、既存ボリューム全ての一覧を取得し、全てのボリュームについて要求アクセスタイムを満たせるかどうかを調べたが、これを指定されたパリティグループと同じ物理ディ

スクを使用している全てのボリュームもしくは指定したパリティグループ内の全てのボリュームのみについてしらべてもよい。

【0028】図6は、図4に示す代替案探索ステップ4007の処理手順である。代替案を探索する際は、まずステップ6001で指定された容量と指定されたアクセスタ임을満たせる空き領域を全ての物理ディスクの中から探索する。ここで使用するボリューム構成情報は、先に図4で述べたボリューム構成情報取得ステップ4002で取得した情報を再利用する。ステップ6001の探索の結果、要求された容量と要求アクセスタ임을満たせる空き領域があれば（ステップ6002）、ステップ6007で結果を代替案として終了する。しかしステップ6002で要求を満たせる空き領域がないと判断したときは、ステップ6003で、要求された容量を10%減らした容量について要求アクセスタ임을満たせる空き領域を探索する。さらにその結果に関わらず、ステップ6004で、指定された容量で指定された要求アクセスタ임을10%増やした値（10%長くアクセスに時間がかかってよいことを示す）を要求アクセスタimeとしてこれを満たす空き領域を探索する。そしてステップ6005で、ステップ6003及びステップ6004の探索結果が空かどうかを調べ、その探索結果が空でなければ、ステップ6007で前記ステップ6003とステップ6004の探索結果を代替案として終了する。しかしステップ6005で探索結果が空であると判断されれば、ステップ6006で代替案はないと判断して探索を終了する。以上図6で述べたステップは全て図3の代替案探索部3004で処理する。

【0029】図6の説明において、代替案探索方法としてステップ6003、ステップ6004でそれぞれ要求された容量を10%減らした値、要求アクセスタimeを10%増やした値を使用した、この割合に制限はない。別の方法として本発明の利用者が指定した値を使用したり、過去の探索履歴から代替案が見つかる確率が高い割合を使用したりすることもできる。また代替案が見つからなかった場合、ステップ6003又はステップ6004の割合を変えて、ステップ6003以降を再び実行してもよい。またステップ6003もしくはステップ6004のどちらか一方のみを実行したり、ステップ6003もしくはステップ6004の処理手順を繰り返したりしてもよい。

【0030】図7は、図6のステップ6001、ステップ6003及びステップ6004で使用した、指定した容量又は指定した要求アクセスタimeを満たせる空き領域を探索するステップの処理手順である。まずステップ7001で全てのパリティグループ内の空き領域一覧を取得する。これは図3のボリューム構成情報取得部3008を使用する。次にこの空き領域一覧のうち、指定された容量で指定されたアクセスタimeを満たすかどうか

を調べていない空き領域があるかどうかをステップ7002で判断し、なければ終了する。もしステップ7002で調べていない空き領域があれば、ステップ7003でそのうちの1個を選択する。そしてステップ7004で、当該空き領域にボリュームを作成したと仮定した場合、当該空き領域を構成するパリティグループ内に定義されている全てのボリュームが、ボリューム作成後も要求アクセスタimeを満たせるかどうか、図5に示した処理手順で調査する。次にステップ7005で、ステップ7004の結果、要求アクセスタimeが満たせないと判断すれば終了する。ステップ7005でボリューム作成後も要求アクセスタimeを満たせると判断すれば、ステップ7006で空き領域を探索結果に加え、再びステップ7002に戻り、処理を継続する。なお図7に示したステップのうち、ステップ7004については図3の要求アクセスタime維持確認部3005で処理し、それ以外のステップは全て代替案探索部3004で処理する。

【0031】図8は、図5におけるボリューム作成後アクセスタime決定ステップ5003の処理の手順である。まずステップ8001で、ボリューム一覧に含まれる全ての既登録ボリュームの平均I/O回数の合計を計算する。次にステップ8002で、ボリューム作成要求時に指定された平均I/O数をステップ8001の結果に足す。ステップ8002の結果により、要求されたボリュームを定義した後の当該パリティグループ内に定義されているボリュームの平均I/O数の総和が求められる。ステップ8002の処理終了後、ステップ8003でボリューム作成後アクセスタime決定テーブルを参照し、その結果をボリューム作成後アクセスタimeとする。ただしステップ8003で、ボリューム作成後アクセスタime決定テーブルの表の値が式である場合は、当該式を計算し、計算結果をボリューム作成後アクセスタimeとする。

【0032】本実施例ではボリューム作成後アクセスタime決定ステップ5003を、図3における管理計算機2004内のボリューム作成後アクセスタime決定部3009において実現する例を示したが、本ステップを実行するブロックをストレージ3011内に設けてもよい。この場合、ストレージ3011内に新規にパリティグループの性能履歴を採取するパリティグループ性能履歴採取部と、ボリューム作成後アクセスタime決定部と、ボリューム作成後アクセスタime決定要求受け付け部を設け、このボリューム作成後アクセスタime決定部は、パリティグループ性能履歴採取部が採取した履歴からボリューム作成後アクセスタime決定要求受け付け部が受け取った要求により指定された平均I/O数に最も近い履歴の平均アクセスタimeを抽出し、その抽出結果をボリューム作成後アクセスタimeとする。このとき管理計算機2004では、ボリューム作成後アクセスタime

ム決定部3009の代わりにボリューム作成後アクセス
タイム決定要求送信部を新規に設け、ストレージ301
1内のアクセスタイムボリューム作成後アクセスタイム
決定部に要求を発行する。

【0033】図9は、図8に示したボリューム作成後ア
クセスタイム決定ステップのステップ8003で参照す
るボリューム作成後アクセスタイム決定テーブルを示し
たものである。列9001はボリューム作成要求と共に
指定されたボリューム定義先のパリティグループ内に定
義されているボリュームの平均I/O数の合計であり、
列9002はその平均I/O数の合計に対して当該パ
リティグループが達成可能な平均アクセスタイムの値も
しくは式を示す。なお図9に示したボリューム作成後ア
クセスタイム決定テーブルは、ストレージ内での全てのパ
リティグループで共通のテーブルを持っても良いし、パ
リティグループ毎にもってもよい。

【0034】図9に示した数値は例であり、実際の場
合は異なる値を使用してもよい。また常に固定した表を
使用するのではなく、図2に示すクライアント計算機20
02やクライアント計算機2003上で性能履歴を収集
し、当該パリティグループもしくは構成する物理ディス
クの構成及び性能が同じ他のパリティグループの平均I
/O数と平均アクセスタイムを抽出し、その結果を管理
計算機2004に送って、ボリューム作成後アクセス
タイム決定テーブルに反映してもよい。

【0035】またアクセスタイム決定テーブルの値を当
該パリティグループを構成する物理ディスクの台数やそ
れらの性能、カタログに記述されている仕様性能値、当
該パリティグループのRAIDレベルや利用可能なキャ
ッシュメモリの容量から計算してもよい。例えば平均ア
クセスタイムを求めるのに、当該パリティグループを構
成する物理ディスクの平均アクセスタイムの平均値も
しくは最も遅い値にパリティグループのRAIDレベルに
より決定される固定値を積算する等である。このときの
RAIDレベルにより決定される固定値の例として、R
AIDレベル1の場合は2倍、データ用物理ディスクを
3台と冗長データ1台を使用するRAIDレベル5の場
合は1/2等がある。

【0036】図10は、図3に示す要求アクセスタイム
保持部3006が保持する情報である。列10001
は、ボリュームの識別子を示し、列10002及び列1
0003は、それぞれ列10001で示されたボリューム
の要求アクセスタイムと平均I/O数を示す。

【0037】本実施例では、要求アクセスタイムを図3
の要求アクセスタイム保持部3006で集中管理する例
を述べたが、クライアント計算機、管理計算機もしくは
ストレージの内にまとめて保持したり、ストレージ内の
ボリューム定義ごとにもったりしてもよい。

【0038】図11は、図3に示すストレージ3011
内のボリューム情報保持部3014が保持するボリュー

ム構成情報を示す。図11(a)は、パリティグループ
と当該パリティグループを構成している物理ディス
クと、そのパリティグループのRAIDレベルの一覧であ
る。列11001はパリティグループの識別子であり、
列11002及び列11003は、それぞれ当該パ
リティグループを構成する物理ディスク一覧及び当該パ
リティグループのRAIDレベルである。また図11(b)
は、パリティグループ内に定義されているボリュームも
しくは空き領域の一覧とそれらの容量を示している。列
11004はパリティグループで、列11005はその
パリティグループ内に定義されているボリューム識別子
もしくは空き領域を示す情報であり、列11006は列
11005で示した内容に割り当てられている容量であ
る。

【0039】本実施例では、指定したボリューム作成
後、要求アクセスタイムを下回るボリュームがでる場合
に警告及び代替案を表示する例を示したが、警告のみで
とどめ代替案を表示しなくてもよい。このときは、図4
においてステップ4007の代替案探索の処理は行なわ
ず、ステップ4008では警告のみを表示する。また警
告を出した後、要求されたボリュームを作成するかどうか
を確認し、継続すると選択されたら、他のボリューム
のアクセスタイムが要求アクセスタイムを満たさなくな
ってもボリュームを作成するようにしてもよい。このと
き図4の処理の流れは、ステップ4008の後、継続確
認をして継続を選択したらステップ4006を、中止を
選択したら終了するという流れになる。

【0040】また本実施例では、ボリュームの作成に限
定して述べたが、ボリュームの移動に関しても同様であ
る。ボリュームの作成の場合は、ボリュームを作成する
パリティグループに着目して性能に関する影響を考慮し
たが、ボリュームの移動の場合は、ボリュームの移動先
のパリティグループに着目して性能に関する影響を考慮
する。

【0041】図9で示したボリューム作成後アクセス
タイム決定テーブル及び図10で示した要求アクセス
タイム保持部3006が保持する情報は、管理者が指定し
て変更する以外に、指定した時間に自動的に変更しても
よい。例えばタイマ部を設けて、テーブルエントリを定
期的に変更したり、有効期限をつけたりするなどである。
また図3で示した要求アクセスタイム保持部3006が
保持する情報について、要求アクセスタイムに優先度
をつけ、優先度により図4のステップ4008で警告を表
示して終了するか、前述のように継続確認するかを切り
替えてもよい。

【0042】次に第2の実施例として、要求アクセス
タイムを指定し、その要求を満たす空き領域を表示する
手順を示す。図12は、要求アクセスタイムを指定し、
それを満たせる空き領域の一覧を表示する手順である。ま
ずステップ12001で要求アクセスタイムを取得す

る。次にステップ12002で、その要求アクセスタイムと容量500メガバイト以上を条件として、条件を満たす空き領域を探索する。このときの探索方式は、図6で示した手順を使用する。また容量500メガバイトという条件は特にこの値に限定しない。ステップ12002の次にステップ12003で、探索結果の各空き領域の容量とボリューム定義後アクセスタイムを全て表示する。本実施例によれば、新しいボリュームを作成する前に要求アクセスタイムを満たす空き領域が物理ディスク上に存在するかを否か判定することができる。

【0043】本実施例では、要求アクセスタイムを指定してそれを満たす空き領域の一覧を指定したが、同様に容量を指定して、それ以上の容量の空き領域の一覧を表示してもよい。また容量や要求アクセスタイムを指定せずに、全ての空き領域について、容量とボリューム定義後アクセスタイムを表示してもよい。

【0044】次に第3の実施例として、第1の実施例で利用した各ボリュームの要求アクセスタイムを登録する手順を示す。図13は、既存ボリュームの要求アクセスタイムを登録する手順である。まずステップ13001で要求アクセスタイムを登録するボリュームの情報を取得し、ステップ13002で、図8に示した手順でボリューム作成後アクセスタイムを決定する。そしてステップ13003でボリューム作成後アクセスタイム以上の値に制限して、本発明の利用者に要求アクセスタイムとして登録する値を選択させ、ステップ13004で選択された結果を要求アクセスタイムとして登録する。

【0045】第1の実施例、第2の実施例及び第3の実施例では性能評価基準の例としてアクセスタイムを使用した。ボリュームによってデータ転送速度が異なるならば転送速度でもよい。またアクセスタイムと転送速度を組み合わせた性能基準でもよい。また読み出し性能と書き込み性能を別々に扱ってもよい。また性能に限らず、RAIDレベルやストレージが持つ機能の限定等の信頼性を表す情報や、他のボリュームとの相対的位置関係、例えば指定した2個のボリュームは同じ物理ディスクを使用しない、等の評価基準を使用してもよい。さらにこれらの基準を2個以上混在して指定し、全てが成り立つもしくはそのうちの1個のみが成り立つという基準で警告するかどうかを決定してもよい。

【0046】第3の実施例で示した要求アクセスタイムの登録は、ボリューム作成直後に限らずボリュームを作成してしばらくたってからでも、ボリュームを作成すると決定してからボリュームを作成するまでの間でもよい。また図4におけるボリューム作成ステップ4006の前もしくは後で処理してもよい。

【0047】第1の実施例及び第2の実施例及び第3の実施例では、各ボリュームの平均利用情報として平均I/O数を使用した。ところが当該ボリュームを使用するクライアント計算機が常に登録された平均I/O数前後

で当該ボリュームを使用するとは限らない。場合により、極端にI/O要求数が増えて何倍にもなり、そのために他のボリュームの性能が低下し要求性能を満たさなくなる可能性がある。クライアント計算機の単位時間あたりの平均I/O数が極端に増加するのを回避する方式として、いくつかの方式が考えられる。そのうちの1個として、クライアント側で発行するI/O要求数を制限する方法である。この場合、現在の平均I/O要求発行数がボリューム作成時に指定した平均I/O数の超えた時点から、I/O要求の発行を自ら遅らせる、もしくは一定時間は要求を発行しない等である。また別の方法として、ストレージ側で各ボリュームに指定された平均I/O数を現在の平均I/O要求処理数が超えたら、しばらくの間、当該ボリュームへのI/O要求の処理の優先度を下げる、もしくは一定時間前期要求は処理しないという方法がある。

【0048】

【発明の効果】本発明により、ボリュームの作成又は移動に際して他ボリュームの要求性能を考慮しながらボリュームの作成・移動操作を行なえるようになる。

【図面の簡単な説明】

【図1】物理ディスクとボリュームの関係を説明する図である。

【図2】実施例のコンピュータシステムの構成を示す図である。

【図3】実施例の管理計算機の構成を示す図である。

【図4】実施例のボリューム作成処理の手順を示すフローチャートである。

【図5】実施例のボリューム作成処理後に全ボリュームが要求アクセスタイムを満たすかを調べる手順を示すフローチャートである。

【図6】実施例の代替案を探索する手順を示すフローチャートである。

【図7】実施例の要求アクセスタイムを満たす空き領域を探索する手順を示すフローチャートである。

【図8】実施例のボリューム作成後アクセスタイムを決定する手順を示すフローチャートである。

【図9】ボリューム作成後アクセスタイム決定テーブルの例を示す図である。

【図10】実施例の要求アクセスタイム保持手段の保持する内容を示す図である。

【図11】実施例のストレージ構成情報保持部が保持する情報を示す図である。

【図12】実施例の要求アクセスタイムを満たすボリューム一覧を取得する手順を示すフローチャートである。

【図13】実施例の要求アクセスタイムを登録する手順を示すフローチャートである。

【符号の説明】

2004：管理計算機、2005、2006、200

7、3011：ストレージ、3001：ボリューム作成

要求受け付け部、3002：警告・代替案表示部、3003：ボリューム作成要求処理部、3004：代替案探索部、3005：要求アクセスタイム維持確認部、30

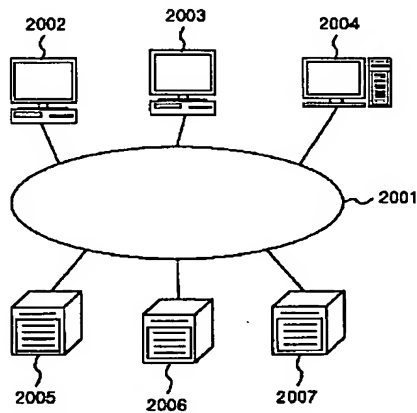
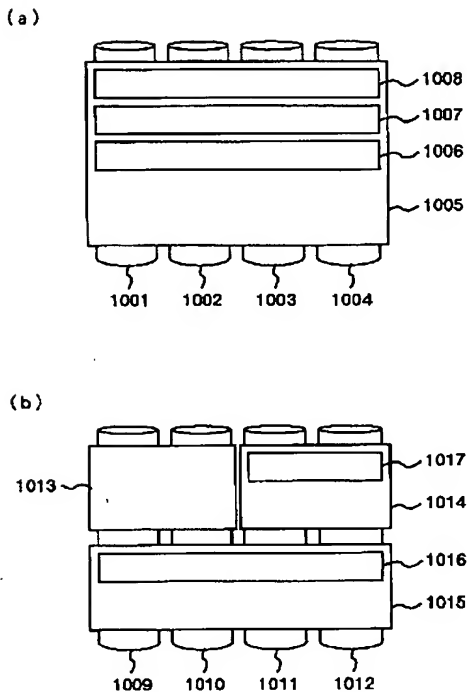
06：要求アクセスタイム保持部、3009：ボリューム作成後アクセスタイム決定部、

【図 1】

【図 2】

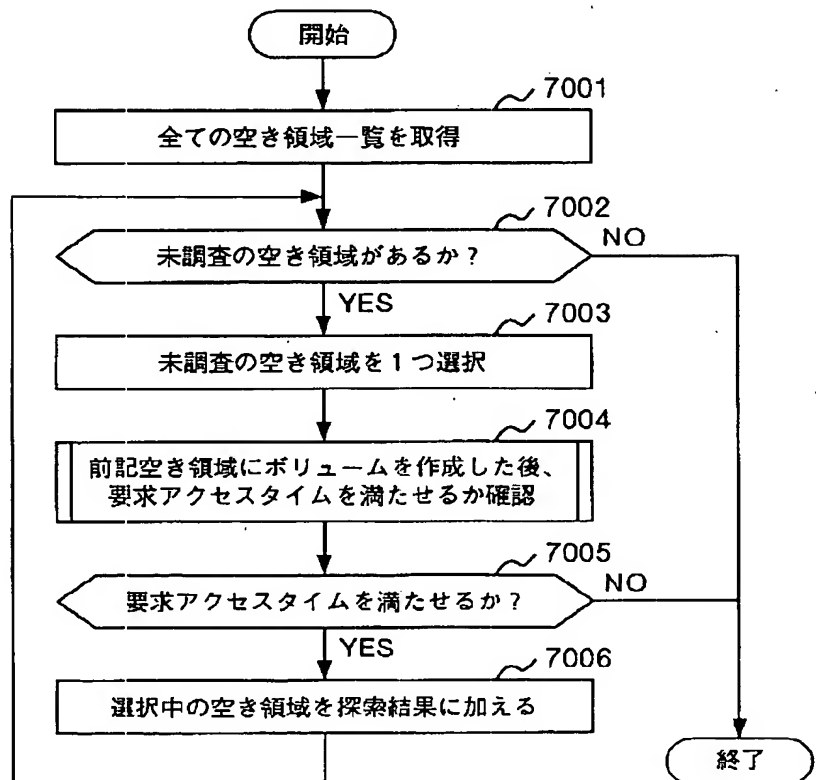
図 1

図 2



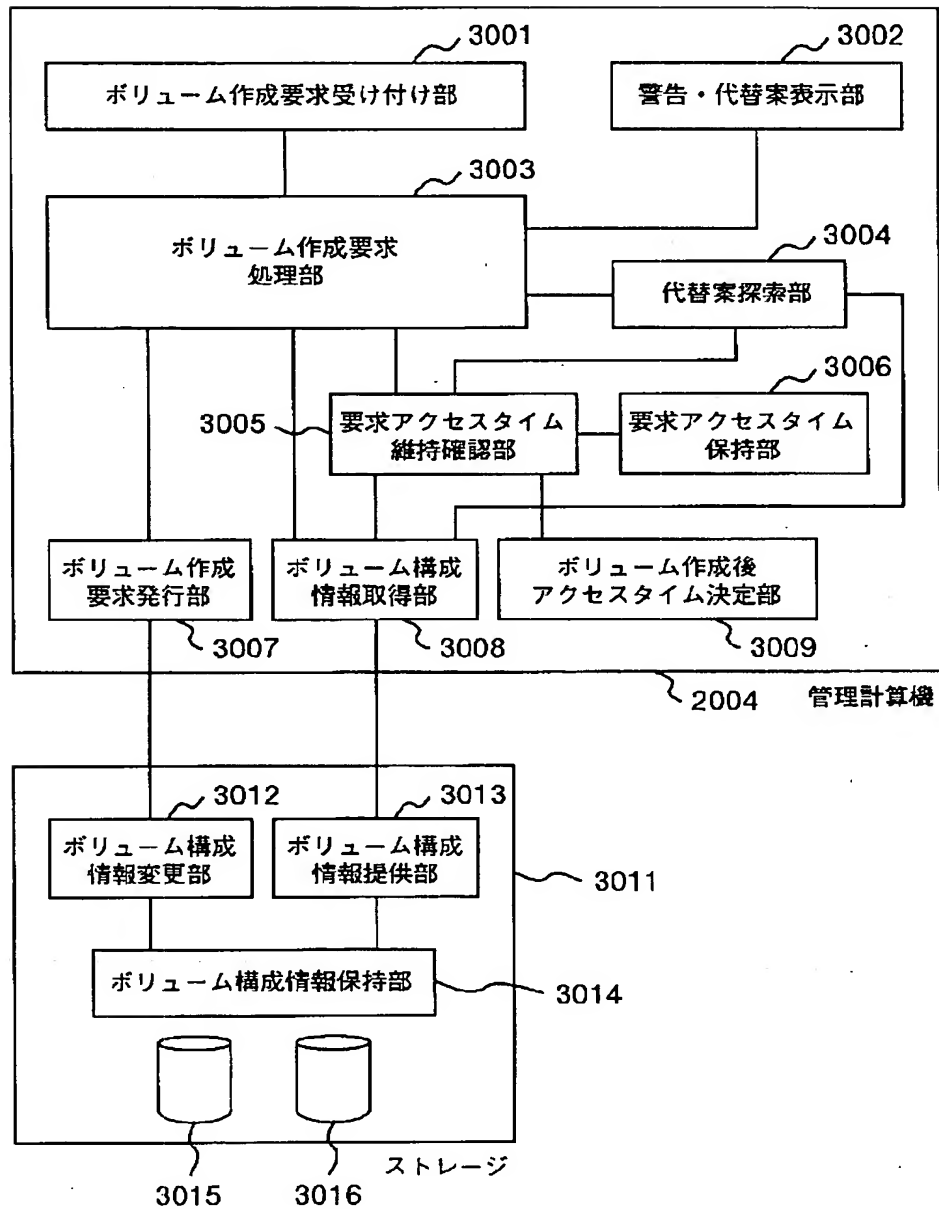
【図 7】

図 7



【図3】

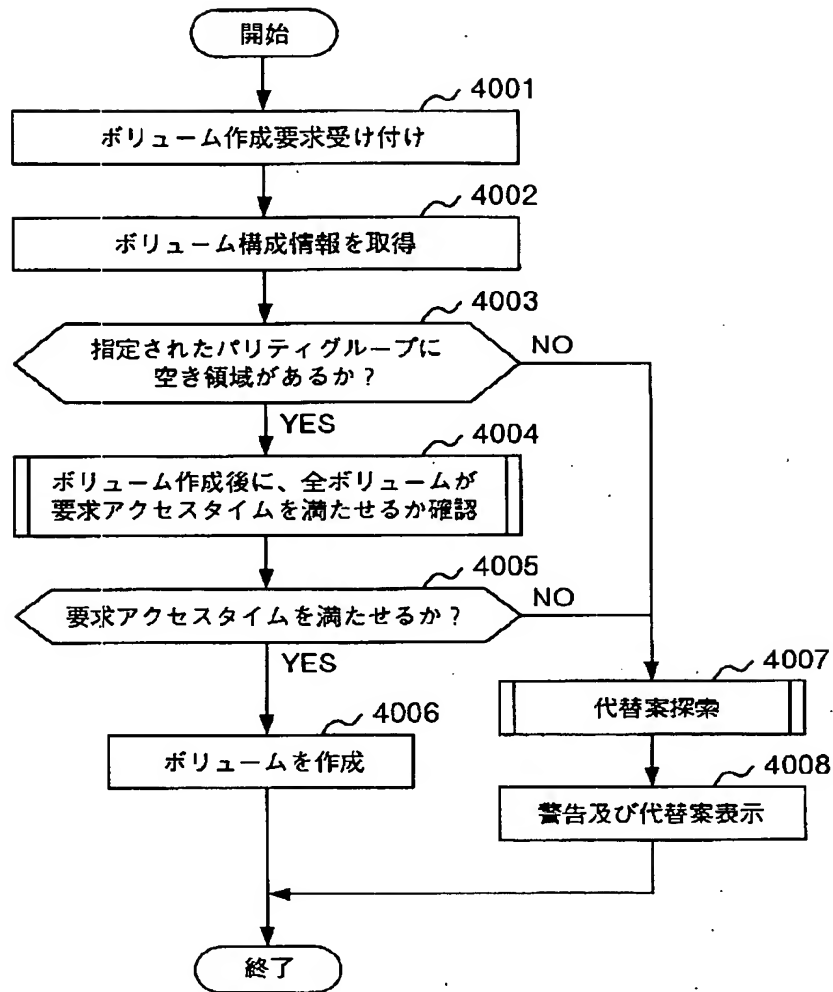
図 3



BEST AVAILABLE COPY

【図4】

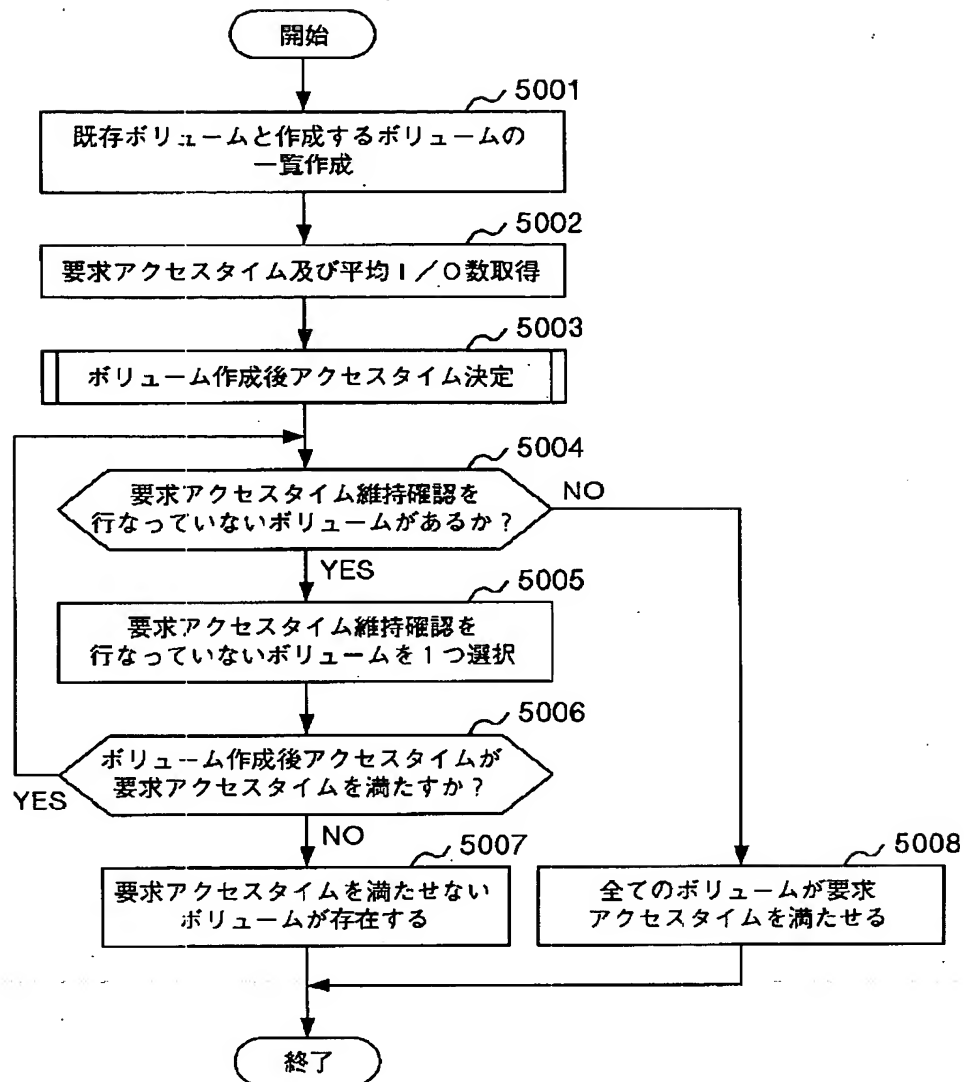
図 4



BEST AVAILABLE COPY

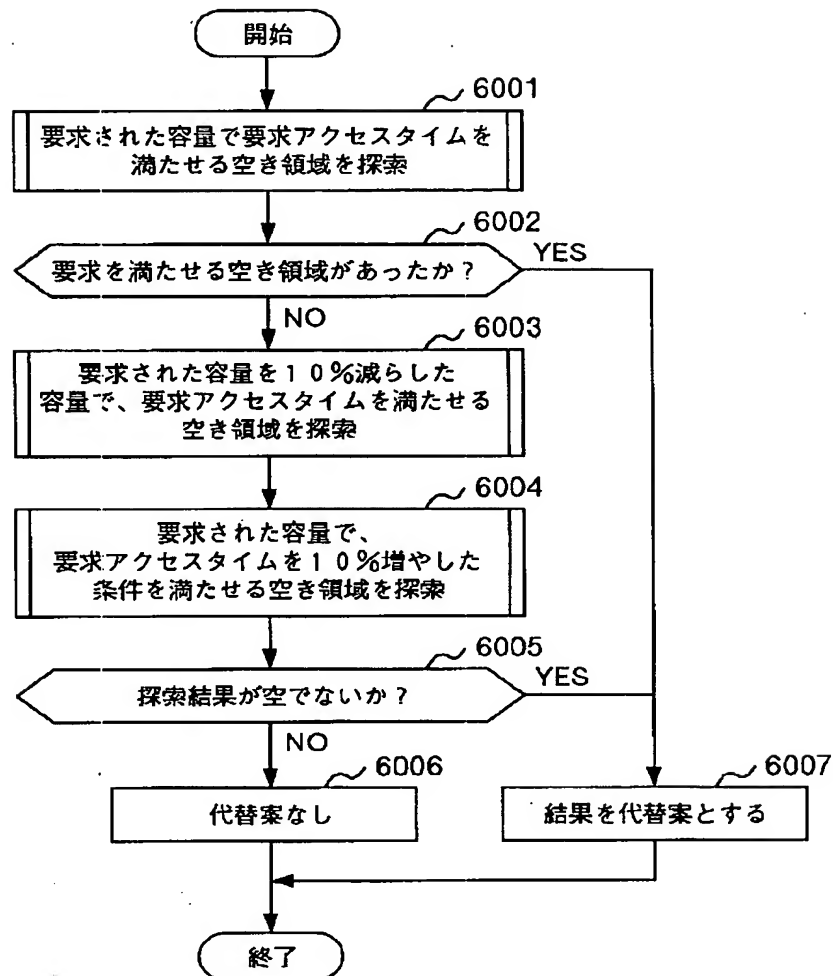
【図5】

図 5



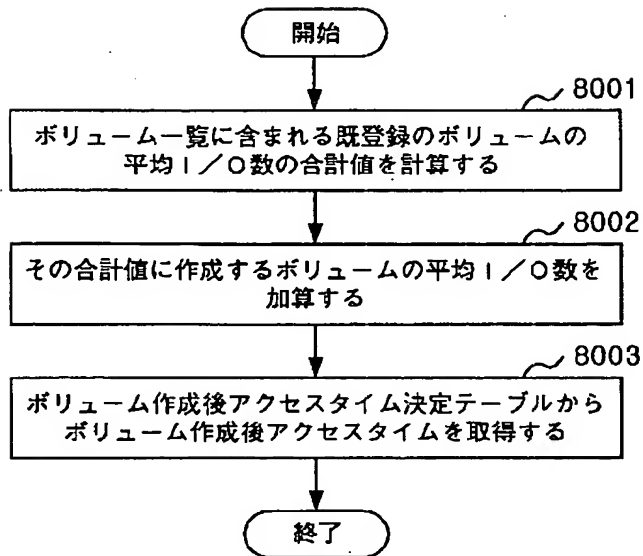
【図6】

図 6



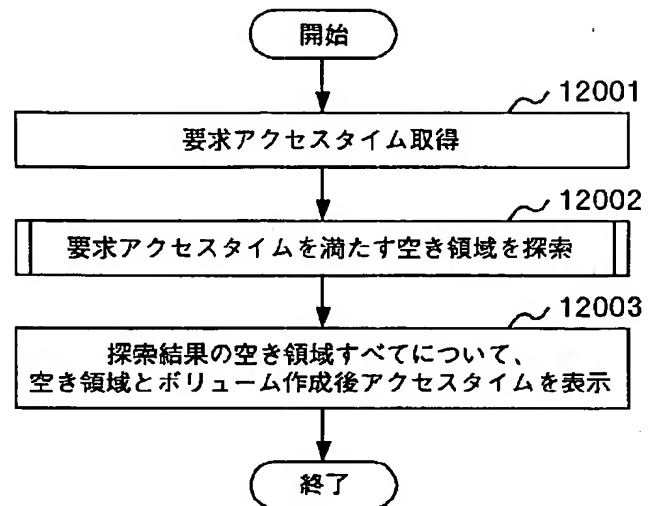
【図8】

図 8



【図12】

図 12



【図9】

図 9

同一パリティグループへの 平均I/O数の合計 (要求/秒)	ボリューム作成後アクセスタイム
0-199	20msec
200-499	24msec
500-799	30msec
800-999	40msec
1000 以上	"平均I/O数の合計" / 10msec

9001

9002

BEST AVAILABLE COPY

【図10】

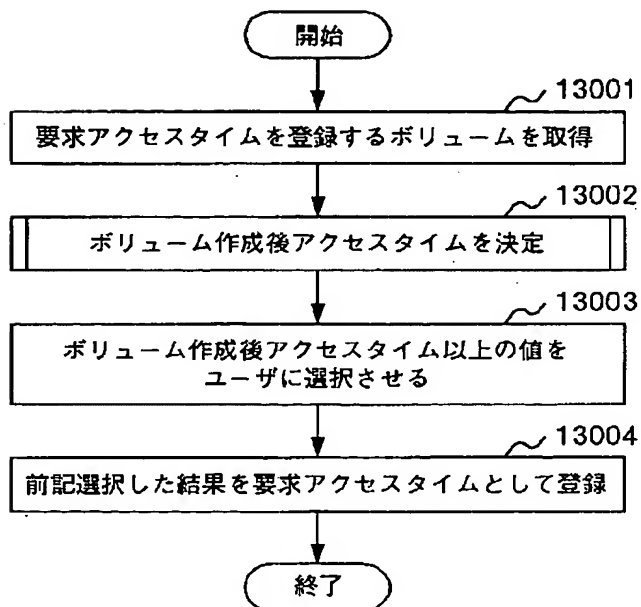
図 10

ボリューム	要求アクセスタイム	平均I/O数(要求/秒)
ボリューム1	30msec	100
ボリューム2	40msec	200
ボリューム3	55msec	150
ボリューム4	45msec	200
ボリューム5	35msec	300

10001 10002 10003

【図13】

図 13



【図11】

図 11

(a)

パリティグループ	物理ディスク	RAIDレベル
パリティグループ1	物理ディスク1、2	RAID-5
パリティグループ2	物理ディスク3～6	RAID-1
パリティグループ3	物理ディスク7～11	RAID-5
パリティグループ4	物理ディスク12～15	RAID-5

}
}
}

11001
11002
11003

(b)

パリティグループ	割り当て内容	容量 (Mbyte)
パリティグループ1	ボリューム1	3000
	空き領域	5000
パリティグループ2	ボリューム2	5000
	ボリューム3	8000
	ボリューム4	2000
	空き領域	5000
パリティグループ3	空き領域	10000
パリティグループ4	ボリューム5	5000
	空き領域	5000

~ 11004
~ 11005
~ 11006

フロントページの続き

(72)発明者 宮崎 聡
 神奈川県小田原市中里322番地2号 株式
 会社日立製作所SANソリューション事業
 部内

Fターム(参考) 5B065 BA01 CA30 CC03 ZA15